

Интеллектуализация сервисов открытого образования

В статье представлен подход к нивелированию основных недостатков систем онлайн-обучения, связанных с трудностями объективного контроля знаний студентов. Предлагается метод интеграции в архитектуру образовательного портала модуля, объединяющего нейросети и генетические алгоритмы с целью более качественной обработки трудноформализуемых и неточных данных. Представлены результаты апробации данного подхода, позволившего поднять сервисы онлайн-обучения на новый уровень автоматизации, гибкости и динамичности.

Ключевые слова: открытое образование, вычислительный интеллект, сервисы электронного обучения.

INTELLECTUALIZATION OF OPEN EDUCATION SERVICES

The article presents the approach of leveling of the main drawbacks of online learning systems associated with difficulties of objective monitoring of students' knowledge. This method is based on the integration of the module that combines neural networks and genetic algorithms in the education system architecture with a goal of better examining of incomplete and inaccurate data. There are described the results of this method, which rises e-learning services to a new level of automation, agility and dynamism.

Keywords: open education, computational intelligence, online services.

Введение

Современное образование позволяет человеку учиться непрерывно, постоянно повышать свой профессиональный уровень. Особую важность приобретает умение находить в огромном объеме данных информацию, необходимую для решения конкретной задачи, и интерпретировать ее согласно своим потребностям. Для профессиональной мобильности и творческой активности человеку необходимо иметь доступ к мировым информационным ресурсам и базам знаний. Поэтому сегодня правильно было бы говорить о современном образовании как о smart-образовании. Smart учебный процесс (образование) – учебный процесс с использованием технологических инноваций и интернета, который предоставляет студентам возможность приобретения профессиональных компетенций на основе системного многомерного видения и изучения

дисциплин, с учетом их многоаспектности и непрерывного обновления содержания [1].

Во всем мире набирают популярность порталы, нацеленные на дистанционное обучение – от краткосрочных тренингов до высшего образования. Динамичное развитие рынка онлайн-образования ужесточает установившуюся на нем конкуренцию. Но в сравнении с традиционным подходом такие неоспоримые достоинства онлайн-обучения, как доступность и удобство, заметно снижаются из-за недостатков, основные из которых заключаются в следующем:

- нехватка специалистов для составления курсов, сведение к минимуму обратной связи с учениками и индивидуального преподавательского подхода;
- сложность объективного контроля знаний. Индивидуальную проверку заданий заменяют различные формы тестирования, при этом отсутствуют рекомендации по

изучению материалов на основании результатов контроля;

- сложность объективного оценивания качества обучения. До сих пор не существует универсального определения показателей качества содержания курсов. Из-за отсутствия критериев количественного оценивания характеристик материалов в курс зачастую включаются устаревшие и недостоверные источники. Программа по дисциплине фиксирована либо варьируется самим пользователем.

Один из путей устранения этих недостатков базируется на последних разработках в области искусственного интеллекта. Гибридизация методов интеллектуальной обработки информации – девиз последних лет в области технологий искусственного интеллекта. Сформировалась концепция «мягких вычислений» (soft computing), объединяющая нечеткую логику, нейросети, эволюционные алгоритмы и вероятностные рассуждения.



Нина Викторовна Комлева,
к.э.н., профессор кафедры
математического обеспечения
информационных систем и
инноватики МЭСИ
Тел.: (495) 442-80-98
Эл. почта: nkomleva@mesi.ru
Московский государственный
университет экономики, статистики
и информатики
www.mesi.ru

Nina V. Komleva,
PhD in Economics, Professor,
Department of Information Systems
Software and Innovation, MESI
Tel.: (495) 442-80-98
E-mail: nkomleva@mesi.ru
Moscow State University of Economics,
Statistics and Informatics
www.mesi.ru

В частности, методы нечеткой логики в совокупности с концепцией нейросетей зарекомендовали себя одним из самых эффективных и перспективных инструментов мягких вычислений. Их популярность относительно методов классических разделов математики обусловлена тем, что для решения задач интеллектуальной обработки и анализа данных успешно применяются нейро-нечеткие модели.

1. Основные тенденции развития вычислительного интеллекта

Исследования в области вычислительного интеллекта как инструмента интеллектуализации обработки данных с каждым годом приобретают все большую практическую ценность.

Вычислительный интеллект (computational intelligence, CI) – ответвление искусственного интеллекта и альтернатива его классическим подходам, основанным на строгом логическом выводе. В современном понимании искусственный интеллект (artificial intelligence, AI) – это совокупность методов и инструментов решения сложных прикладных задач, использующих принципы, заложенные в процессах, протекающих в живой или неживой природе (например, в мышлении). CI опирается на эвристические, неточные и приближенные алгоритмы решения задач, зачастую не имеющих решения за полиномиальное время. Он сочетает в себе методы обучения, адаптации, эволюции и нечеткой логики для создания в известной степени интеллектуальных программ.

Таким образом, вычислительный интеллект работает с классом задач из области искусственного интеллекта, но классический AI решает их методами традиционных жестких вычислений, а CI – мягких вычислений.

Мягкие вычисления (soft computing) – это симбиоз вычислительных методологий, коллективно обеспечивающих базу для понимания, конструирования и развития интеллектуальных систем. По срав-

нению с жесткими вычислениями, мягкие вычисления более приспособлены к обработке неточных, неопределенных и неполных данных для достижения удобства манипулирования, робастности, низкой стоимости решения и лучшего согласия с реальностью. Инструментарий CI основан на следующих наиболее активно развивающихся подходах и методах мягких вычислений: нечеткие системы (нечеткие множества, нечеткая логика, нечеткие регуляторы), нейронные сети, эволюционное моделирование (генетическое программирование, эволюционное программирование, генетические алгоритмы), теория хаоса, роевой интеллект, искусственные иммунные системы, гибридные системы (нечеткие нейросети, комбинации генетических алгоритмов и нейросетей).

Вычислительный интеллект дает новый взгляд на статистические методы, примером чему служат нечеткие системы. Нечеткая логика лежит в основе методов работы с неточностью, приближенных рассуждений и вычислений со словами (computing with words). Искусственные нейронные сети, в свою очередь, тесно связаны с машинным обучением. Нейровычисления отражают способность к обучению, адаптации и идентификации. В случае генетических вычислений и роевого интеллекта речь идет о возможности систематизировать случайный поиск и достигать оптимального значения характеристик объекта. Вероятностные вычисления обеспечивают базу для управления неопределенностью и проведения рассуждений, исходящих из свидетельств.

Основой развития вычислительного интеллекта является разработка алгоритмических моделей для решения сложных задач высокой размерности путем моделирования естественных природных систем. Каждая из парадигм CI основана на моделировании реально существующих биологических явлений и свойств. Нейросети зародились как модель фрагментов нервной системы; эволюционные вычисления имитируют естественную эволюцию и используют за-



Ольга Андреевна Хлопкова,
аспирант кафедры математического
обеспечения информационных систем
и инноватики МЭСИ
Тел.: (495) 442-80-98
Эл. почта: perikrone@gmail.com
Московский государственный
университет экономики, статистики
и информатики
www.mesi.ru

Olga A. Khlopkova,
Post-graduate Student of the Department
of Software Information Systems and
Innovation, MESI
Tel.: (495) 442-80-98
E-mail: perikrone@gmail.com
Moscow State University of Economics,
Statistics and Informatics
www.mesi.ru

коны генетики; роевой интеллект моделирует социальное поведение колоний организмов; искусственные иммунные системы не были бы построены без биологических прототипов; нечеткие системы возникли в результате исследования взаимодействия организмов с окружающей средой.

Итак, вычислительный интеллект включает в себя следующие основные парадигмы:

- искусственные нейронные сети (нейросети, artificial neural networks);
- нечеткие системы (fuzzy systems);
- эволюционное моделирование (evolutionary modelling);
- роевой интеллект (swarm intelligence).

Кроме того, в рамках вычислительного интеллекта ведутся исследования искусственных иммунных систем (artificial immune systems), теории хаоса (chaos theory), фрактальных преобразований и вейвлетов.

Нейросеть – это распределенный параллельный процессор, состоящий из элементарных единиц обработки информации – нейронов, накапливающих экспериментальные знания.

Нечеткая система – это система с нечеткими: спецификацией параметров, описанием входных и выходных переменных и выводом на основе базы знаний, состоящей из нечетких продукционных правил вида IF (антецедент) THEN (консеквент). Нечеткая система состоит из базы знаний, фаззификатора, функций термов, ядра нечеткого логического вывода и в большинстве случаев дефаззификатора.

Как показывает практика, нечеткое управление дает лучшие результаты в сравнении с получаемыми при помощи классических алгоритмов. Очевидной областью внедрения нечеткой логики являются всевозможные экспертные системы, в том числе: нелинейный контроль и управление в производстве, самообучающиеся системы, системы, распознающие тексты на естественном языке, системы планирования и прогнозирования, опирающиеся на неполную инфор-

мацию, финансовый анализ в условиях неопределенности, управление базами данных, совершенствование стратегий управления и координации действий.

Эволюционное моделирование (ЭМ) – это класс вычислительных методов и алгоритмов поиска, оптимизации и обучения. Концепция эволюционного моделирования основывается на принципах естественного отбора (особи, которые лучше способны решать задачи в своей среде, выживают и репродуцируют), скрещивания (хромосома потомка состоит из частей хромосом предков), мутации (приобретение потомком свойств, отсутствующих у родителей).

К методам ЭМ относятся эволюционные стратегии, эволюционное программирование, генетические алгоритмы и генетическое программирование. Их отличия в основном состоят в способе представления решений и наборе эволюционных операторов.

Генетические алгоритмы (ГА) – оптимизационные методы, основанные на использовании аналогий с процессами естественного отбора и генетических преобразований.

Важным достоинством ГА является его способность манипулировать одновременно большим множеством разнородных параметров-признаков, из которых комбинируется хромосома. ГА обладают высокой помехозащищенностью и робастностью: при снижении надежности входных данных качество их работы снижается очень медленно.

Эволюционные стратегии тоже относятся к эвристическим методам оптимизации, основанным на адаптации и эволюции, но отличаются от ГА реализацией таких шагов алгоритма, как отбор и мутация, а также способом представления генотипа.

Генетическое программирование – автоматическая генерация программ с помощью генетических алгоритмов. Одним из основных вопросов генетического программирования является задание способа кодирования, оптимального для внесения мутации и объединения алгоритмов оператором скрещива-

ния. Способы кодирования можно разделить на прямое кодирование для обработки кода программы в явном виде и косвенное кодирование для обработки правил его построения.

Эволюционное программирование аналогично генетическому программированию, но предполагает, что структура программы фиксирована, изменяются только переменные.

ЭМ применяется для автоматизации решения различных оптимизационных задач науки и техники, изучения и моделирования отдельных процессов естественной эволюции, совершенствования искусственных систем за счет наделения их свойствами адаптивного поведения и самоорганизации.

Роевой интеллект (swarm intelligence) описывает коллективное поведение многоагентной децентрализованной самоорганизующейся системы. Методы роевого интеллекта схожи с эволюционными вычислениями и основываются на моделировании социального поведения живых организмов. В задачах численной оптимизации применяются роевые алгоритмы (particle swarm optimization, PSO), в задачах комбинаторной оптимизации – муравьиные алгоритмы (ant colony optimization, ACO). Известны алгоритмы, имитирующие поведение пчел, рыб, птиц и других социальных организмов.

Алгоритмы роевого интеллекта отличаются простотой реализации и могут быть использованы для формирования поведения множества мобильных роботов – агентов системы. На основе данных алгоритмов производится позиционирование агентов в пространстве и прокладывание их оптимальных маршрутов к заданным целям.

Между методами эволюционного моделирования и роевого интеллекта прослеживается сходство. Например, PSO напоминает ЭА тем, что оба подхода основаны на популяциях и вычислении функций приспособленности для каждой особи. Корректировка частиц в PSO схожа с оператором скрещивания, используемым в ЭА.

Тем не менее PSO имитирует именно социальное поведение, а не выживание наиболее приспособленных особей. Другим важным различием является способность частиц извлекать пользу из предыдущего опыта, в то время как ЭА подобных механизмов не предусматривает. Кроме того, по сути роевые алгоритмы являются непрерывными, в то время как ГА в общем случае считаются дискретными.

Методы вычислительного интеллекта чрезвычайно разнообразны и многое позаимствовали из биологии, нейрофизиологии, генетики, социологии и психологии.

Специфика CI состоит в том, что они используют вероятностные, а не детерминированные правила генерации решений, позволяют накапливать и использовать знания о пространстве поиска и, следовательно, проявляют способность к самообучению.

На протяжении всей истории вычислительного интеллекта закономерно прослеживается консолидация его методов. Так, в настоящее время термины «эволюционное программирование», «эволюционные стратегии», «генетические алгоритмы» и «генетическое программирование» рассматриваются как частные ответвления общего подхода «эволюционного моделирования», иногда называемого «эволюционными вычислениями». Нейросетевые системы все чаще интегрируются с алгоритмами нечеткого вывода, генетическими алгоритмами или методами роевого интеллекта. Это позволяет повысить скорость и качество обучения нейросетей, сделать их более динамичными и надежными.

Особенную популярность получили гибриды нейросетей с нечеткими системами и эволюционными вычислениями: нейро-нечеткие регуляторы, COGANN (Combinations of genetic algorithms and neural networks) и т.д. Одним из первых применений роевого алгоритма (particle swarm optimization, PSO) стало именно обучение нейросетей.

В случае гибридизации PSO и генетических алгоритмов (ГА)

чаще всего заимствуется популяция из одного алгоритма и используется в качестве начальной популяции для другого, в частности при снижении скорости роста приспособленности. Сочетание НС и ГА все чаще применяется в предметных областях, где отсутствует физическое или статистическое понимание системы, наблюдается большой статистический разброс входных данных и нелинейность механизма их обобщения.

На данный момент технология последовательных вычислений подошла к пределу своих технических возможностей. Не удивительно, что исследования вычислительного интеллекта переживают очередной подъем: многие методы CI оперируют атомарными объектами для создания гибкой динамичной структуры. Так, нейроны – это примитивные процессоры, которые составляют базу нейросети любой размерности и топологии. Популяции в ГА состоят из особей; роевой интеллект оперирует отдельными частицами (particles), являющимися потенциальным решением задачи.

2. Интеграция COGANN в сервисы онлайн-обучения

Сервисы онлайн-обучения оперируют трудноформализуемыми и неточными данными, среди которых, например, качество и актуальность материалов, корреляция между содержимым различных курсов, текущий уровень знаний по дисциплине, вероятность самостоятельного прохождения теста и т.д. Механизм обобщения таких данных не является линейным, и в качестве инструмента их обработки могут выступать методы COGANN (Combinations of genetic algorithms and neural networks).

Сервисы портала онлайн-образования функционируют в условиях ограниченной и неполной информации, а накапливаемая статистика носит стохастический характер. Концепция COGANN имеет высокий потенциал как средство автоматизации работы такой сложной динамической системы.

Обобщенная архитектура портала онлайн-обучения с интег-

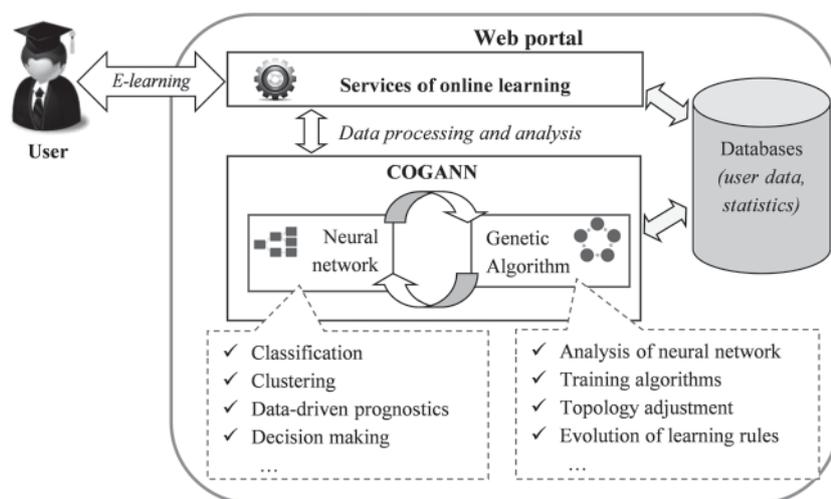


Рис. 1. Модель архитектуры портала онлайн-обучения

рированными в него сервисами COGANN представлена на рис. 1.

Данная модель была реализована и прошла успешные испытания в рамках разработки информационно-аналитической системы НИФИ. Модуль, инкапсулирующий методы COGANN, представлен в виде веб-сервиса. Преимущества сервис-ориентированного подхода (такие как низкая связанность и независимость от платформы) в данном случае играют особую роль, так как позволяют быстро и с минимальными затратами интегрировать модуль в систему.

Как показали практические испытания, использование COGANN привносит адаптивность в онлайн-обучение и предоставляет множество новых возможностей, некоторые из которых:

- анализ статистики по portalу для выявления наиболее востребованной информации, модификации структуры и содержания онлайн-курсов. В рекомендательном порядке формируются и ранжируются по различным критериям списки материалов, подлежащих удалению, корректировке или перемещению между курсами;

- анализ активности пользователя для составления списков материалов для факультативного изучения, изменения уровня сложности контрольных работ, формирования дополнительных разделов дисциплин, модификации проверочных заданий;

- прогнозирование действий пользователя на основании его пре-

дыдущей активности для составления рекомендаций и оптимизации процесса обучения. Такой подход позволяет повысить эффективность обучения и избавить пользователя от рутинных действий;

- динамическое управление процессом тестирования. Если пользователь дает быстрые и правильные ответы, имеет смысл повысить сложность заданий или скорректировать их направленность в рамках темы тестирования. Данный метод повышает точность оценивания знаний и определения пробелов в обучении;

- заключение о степени самостоятельности работы пользователя во время тестирования на основании анализа разнообразных статистических данных.

Перечисленные задачи, сводящиеся к кластеризации, прогнозированию и управлению, эффективно решаются с помощью нейросетей различной топологии: самоорганизующихся карт Кохонена, сетей Хопфилда, нечетких нейросетей, гибридных моделей и т.д.

Например, формирование ряда показателей, косвенно свидетельствующих о самостоятельности прохождения контрольных заданий, реализуется с применением соревновательных нейросетей с обучением без учителя. В то время как формирование рейтинга курсов и учебных материалов может осуществляться нечеткой нейросетью TSK (Takagi-Sugeno-Kanga), для

обучения которой принято использовать гибридный алгоритм.

Обширная статистика развивающихся порталов подходит для них в качестве обучающей выборки и позволяет избежать временных затрат на ручное составление данных. Но для эффективности обучения нейросетей необходимо, чтобы этот процесс был простым и доступным для администраторов портала.

Он автоматизируется за счет применения генетических алгоритмов. Кроме того, даже располагая статистической информацией, необходимо произвести ее предобработку для адекватного обучения и функционирования нейросети. Генетические алгоритмы позволяют выбрать доминантные признаки для формирования полной и непротиворечивой обучающей выборки, которые оказывают наибольшее влияние на вывод НС.

3. Результаты эксперимента

Применение в совокупности с нейросетями генетических алгоритмов не только повышает представительность обучающего множества и эффективность нейросетевого вывода. Генетические алгоритмы также нивелируют ряд закономерных в рамках парадигмы нейросетей проблем: локального оптимума, «паралича» нейросети, переобучения, скорости сходимости итеративных градиентных алгоритмов обучения и т.д.

В реальных условиях эксплуатации параметры портала как исследуемой среды постепенно изменяются во времени. Но динамическая структура входных данных накладывает требование динамичности и на модель обработки: нейросеть должна подстраивать архитектуру и топологию под внешние условия.

Генетические алгоритмы, общие для нейросетей различных архитектур, имеют преимущество как наиболее универсальный способ обучения. Для их использования не требуется информация о характеристиках функции, экстремум которой необходимо найти. При решении практической задачи оптимизации работы сервисов он-

лайн-обучения эти сведения не известны, и вместо них доступны статистические или стохастические данные, хранящиеся на портале.

Генетические алгоритмы уточняют параметры и топологию нейросети на этапе функционирования, повышают надежность ее вывода, при необходимости оптимизируют правила обучения. Следует подчеркнуть, что они способны значительно расширить функциональность нейросети привнесением в ее работу большей гибкости и динамичности.

Заключение

История вычислительного интеллекта как ответвления AI насчитывает более полувека. За это время интерес к СИ волнообразно спадал и нарастал, что можно проследить по этапам его становления, от зарождения в 1950-х гг. вплоть до современности. К сегодняшнему дню СИ считается самостоятельным научным направлением. Он накопил богатый арсенал инструментов, а его теоретическая база и области практического применения продолжают интенсивно расширяться.

Методология вычислительного интеллекта представляет совокупность принципов, подходов, концепций и конкретных методов, а также

методик. Исследования СИ ведутся по следующим наиболее значимым направлениям: нечеткие системы (нечеткие множества, нечеткая логика, нечеткие регуляторы), нейронные сети, эволюционное моделирование (генетическое программирование, эволюционное программирование, генетические алгоритмы), теория хаоса, роевой интеллект, искусственные иммунные системы, гибридные системы (нечеткие нейросети, комбинации генетических алгоритмов и нейросетей).

Как и в любой науке, многие области исследований СИ пересекаются. Взаимосвязь между различными направлениями СИ выражена особенно сильно и породила новое направление: гибридизацию методов вычислительного интеллекта.

Гибридизация вместе с вовлечением в исследования многих предметных областей, имеющих скорее практическое отношение к AI, чем фундаментальное, наделяют СИ системы уникальными свойствами. К ним относятся адаптивность, обучаемость, робастность и универсальность.

Ежегодно возрастающий интерес к вычислительному интеллекту объясняется множеством актуальных областей его применения, среди которых распознавание речи и образов, машинный перевод, авто-

матическое доказательство теорем, аппроксимация функционалов, прогнозирование на биржевом рынке, управление бизнес-процессами, построение самообучающихся систем поддержки принятия решений.

Разумное и обоснованное использование вариаций сочетания нейросетей и генетических алгоритмов в целом повышает технический уровень и качество вычислительных работ. Как подтвердила практическая эксплуатация, системы онлайн-обучения являются целесообразной областью применения методов COGANN.

Интеллектуальные модули портала электронного обучения снижают необходимость вмешательства эксперта в процесс обучения, автоматизируют функции, традиционно возлагающиеся на преподавателя. Они избавляют администраторов от рутинной и трудоемкой работы, связанной с обработкой статистики и корректировкой проверочных заданий. Благодаря внедрению таких модулей заметно сокращаются затраты ресурсов на принятие решения об обработке контента.

Интеграция COGANN в сервисы онлайн-обучения открывает новые перспективы в совершенствовании процесса дистанционного образования и выводит его на новый уровень интеллектуализации.

Литература

1. Россия на пути к Smart-обществу: монография / под ред. проф. Н.В. Тихомировой, проф. В.П. Тихомирова. – М.: НП «Центр развития современных образовательных технологий», 2012.
2. Тельнов Ю.Ф., Третьяков В.М. Интеллектуальные информационные системы. – М.: МЭСИ, 2010. – 222 с.
3. Рутковская Д., Пилиньский М., Рутковский Л. Нейронные сети, генетические алгоритмы и нечеткие системы. – М: Горячая Линия – Телеком, 2008. – 452 с.
4. Емельянов С. Искусственный интеллект и принятие решений. Выпуск 1. – М.: ИСА РАН, 2013. – 88 с.
5. Горяшко А.П., Комлева Н.В. Образование на основе компетенций в открытых информационных средах: алгоритмы принятия решений // Высшее образование в России. Научно-педагогический журнал Министерства образования и науки РФ. – 2011. – № 8–9. – С. 78–84.
6. Комлева Н.В. Модели и инструменты инновационного развития образования в открытой информационной среде: монография. – М.: МЭСИ, 2013. – 199 с.
7. Комлева Н.В., Хлопкова О.А. Обработка контента в информационных средах на основе нейро-нечеткой модели принятия решений // Экономика, статистика, информатика. Вестник УМО. Научно-практический журнал. – 2013. – № 5. – С. 188–192.
8. Комлева Н.В. Моделирование процесса создания открытых электронных образовательных ресурсов // Открытое образование. Научно-практический журнал. – 2013. – № 6. – С. 24–30.
9. Комлева Н.В. Интеллектуальные технологии в совершенствовании научно-исследовательской работы в системе научных институтов и центров финансового сектора // Научно-исследовательский финансовый институт. Финансовый журнал. Научно-практическое издание. – 2013. – № 3. – С. 145–150.
10. Комлева Н.В., Хлопкова О.А. Технология мягких вычислений в автоматизации обработки данных на примере информационно-аналитической системы НИФИ Минфина России // Ценности и интересы современного

- общества (секция «Современные парадигмы информационных технологий в развитии общества»): материалы Международной научно-практической конференции 29 октября 2013 г. – М.: МЭСИ, 2013.
11. *Комлева Н.В.* Открытые образовательные ресурсы, их адаптация и повторное использование // Информационные технологии в образовательном процессе современного университета: теоретические и методические аспекты / под ред. к.т.н. С.В. Федосеева, к.э.н. Н.В. Комлевой. – М.: МЭСИ, 2014. – С. 33–54.
12. *Andreichicov, A.V., Andreichicova, O.N.* New Paradigms of Decision-Making. – Halifax, Canada: Management Science Division, 2003. – P. 28–34.
13. *Haykin S.* Neural networks and learning machines. – Hamilton, Ontario, Canada: Pearson Prentice Hall, 2009.
14. *Holland J.* Evolutionary computation: Building blocks, cohort genetic algorithms, and hyperplane-defined functions. Massachusetts Institute of technology, 2000.
15. *Komleva N.* Innovative information environment for enriching education quality // Int. J. Foresight and Innovation Policy. – 2010. – Vol. 6, No. 4.